

The power of linear programming for valued CSPs: a constructive characterization

Vladimir Kolmogorov
Institute of Science and Technology Austria
vnk@ist.ac.at

Abstract

A class of valued constraint satisfaction problems (VCSPs) is characterised by a valued constraint language, a fixed set of cost functions on a finite domain. An instance of the problem is specified by a sum of cost functions from the language with the goal to minimise the sum.

We study which classes of finite-valued languages can be solved exactly by the *basic linear programming relaxation* (BLP). Recently, Thapper and Živný showed [17] that if BLP solves the language then the language admits a *binary commutative fractional polymorphism*. We prove that the converse is also true. This leads to a necessary and a sufficient condition which can be checked in polynomial time for a given language. In contrast, the previous necessary and sufficient condition due to [17] involved infinitely many inequalities.

1 Introduction

We consider a particular linear programming relaxation of a class of optimization problems called in this paper *basic LP* (BLP). This relaxation has been studied extensively in various domains, especially for objective functions with unary and binary terms. Researchers analyzed its properties, developed efficient algorithms for (approximately) solving this LP, and applied to large-scale practical problems [15, 11, 18, 8, 19, 5, 13, 16, 1, 9].

It has been long recognized that for some classes of optimization problems (e.g. for submodular functions on a chain) BLP relaxation is guaranteed to be tight (i.e. the integrality gap is zero), and allows to solve the problem exactly. A natural question is whether it is possible to characterize such classes.

One possible way to pose the problem formally is to use the framework of *Valued Constraint Satisfaction Problems (VCSPs)* [4]. In this framework a class of allowed objective functions is specified by a *language* Γ , which is a collection of cost functions over a fixed domain D . We say that BLP *solves* $VCSP(\Gamma)$ if the relaxation is tight for all functions that can be expressed as a sum of functions from Γ with overlapping sets of variables.

A major step in the characterization of languages that can be solved by BLP has been recently made by Thapper and Živný [17]. They gave a sufficient condition for BLP to solve $VCSP(\Gamma)$ that covers many known languages such as (1) submodular functions on arbitrary lattices; (2) k -submodular functions; (3) weakly (and hence strongly) tree-submodular on arbitrary trees.

Thapper and Živný also presented a necessary and sufficient condition for BLP to solve $VCSP(\Gamma)$. However, their characterization has one drawback: it involves infinitely many inequalities, which leaves an open question whether checking the condition is a decidable problem for a given finite language Γ .

We resolve this question affirmatively for finite-valued languages Γ . As our **main contribution**, we show that BLP solves such Γ iff Γ admits a *fractional symmetric polymorphism of arity k* for some $k \geq 2$. We prove this result using a mixture of algebraic tools and techniques from Linear Programming.

Related work Kun et al. [12] studied the BLP relaxation for CSP problems and for robust approximability of Max-CSPs. They showed, in particular, that BLP robustly decides a CSP language iff it has *width 1*. Width-1 CSPs were introduced in [7]. A simple characterization of such CSPs was given in [6].

Our work heavily exploits the notion of *fractional polymorphisms* [2]. Fractional polymorphisms is a generalization of *multimorphisms* [4]. It is known that they can characterize all tractable VCSPs [2].

We also mention the work of Raghavendra on analyzing SDP relaxations [14]. In particular, results in Chapters 6 and 7 of [14] imply that the basic SDP relaxation solves languages that admit a *cyclic fractional polymorphism*. This includes languages that satisfy our condition.

2 Background and statement of the results

Let D be a finite domain. A *finite-valued language* Γ is a set of cost functions $f : D^n \rightarrow \mathbb{Q}$ where arity $n \geq 1$ may be different for different functions $f \in \Gamma$. The argument of f is called a *labeling*.

Definition 1. An instance \mathcal{I} of the valued constraint satisfaction problem (VCSP) is a function $D^V \rightarrow \mathbb{Q}$ given by

$$\text{Cost}_{\mathcal{I}}(x) = \sum_{t \in T} f_t(x_{v(t,1)}, \dots, x_{v(t,n_t)})$$

It is specified by a finite set of nodes V , finite set of terms T , cost functions $f_t : D^{n_t} \rightarrow \mathbb{Q}$ of arity n_t and indices $v(t, k) \in V$ for $t \in T, k = 1, \dots, n_t$. A solution to \mathcal{I} is a labeling $x \in D^V$ with the minimum cost. Instance \mathcal{I} is called a Γ -instance if all terms f_t belong to Γ .

The class of optimization problems consisting of all Γ -instances is referred to as $\text{VCSP}(\Gamma)$. Language Γ is called tractable if $\text{VCSP}(\Gamma')$ can be solved in polynomial time for **each** finite $\Gamma' \subseteq \Gamma$. It is called NP-hard if $\text{VCSP}(\Gamma')$ is NP-hard for **some** finite $\Gamma' \subseteq \Gamma$.

One way to tackle a VCSP instance is to formulate and solve a convex relaxation of the problem. Two examples are *basic LP relaxation* and *basic SDP relaxation*, as they are called in [17] and [14] respectively.

The basic LP relaxation will be of particular relevance to this paper. Following [17], we say that basic LP *solves* $\text{VCSP}(\Gamma)$ if for any instance \mathcal{I} from $\text{VCSP}(\Gamma)$ the optimal value of the relaxation equals $\min_x \text{Cost}_{\mathcal{I}}(x)$.

We will study which languages Γ are solved by basic LP. We will do it indirectly by relying on the characterization of [17]. The formulation of basic LP will not be used, and so we omit it in order to avoid unnecessary notation. We refer interested readers to [17] for details.

Fractional polymorphisms We denote \mathcal{O}^m to be the set of operations $g : D^m \rightarrow D$. A *fractional polymorphism* of arity m is a probability distribution ω over \mathcal{O}^m , i.e. a vector with components $\omega(g) \geq 0$ for $g : D^m \rightarrow D$ that sum to 1. Language Γ is said to *admit* ω if every cost function $f \in \Gamma$ of arity n satisfies

$$\sum_{g \in \mathcal{O}^m} \omega(g) f(g(x^1, \dots, x^m)) \leq f^m(x^1, \dots, x^m) \quad \forall x^1, \dots, x^m \in D^n \quad (1)$$

where function f^m is defined via $f^m(x^1, \dots, x^m) = \frac{1}{m}(f(x^1) + \dots + f(x^m))$. We view labelings in D^n as column vectors; given m such columns, operation $g : D^m \rightarrow D$ produces a new column as shown on the right.

x^1	\dots	x^m	$g(x^1, \dots, x^m)$
x_1^1	\dots	x_1^m	$g(x_1^1, \dots, x_1^m)$
\dots	\dots	\dots	\dots
x_n^1	\dots	x_n^m	$g(x_n^1, \dots, x_n^m)$

Operation $g \in \mathcal{O}^m$ is called *symmetric* if it is invariant with respect to any permutation of its arguments: $g(a_1, \dots, a_m) = g(a_{\pi(1)}, \dots, a_{\pi(m)})$ for any

permutation $\pi : [1, m] \rightarrow [1, m]$ and any $(a_1, \dots, a_m) \in D^m$. It is called *cyclic* if $g(a_1, a_2, \dots, a_m) = (a_2, \dots, a_m, a_1)$ for any $(a_1, \dots, a_m) \in D^m$. Note, in the case $m = 2$ both definitions coincide. A fractional polymorphism ω is called symmetric (cyclic) if all operations in $\text{supp}(\omega)$ are symmetric (cyclic). As usual, $\text{supp}(\omega)$ denotes the support of distribution ω : $\text{supp}(\omega) = \{g \in \mathcal{O}^m \mid \omega(g) > 0\}$.

Generalized fractional polymorphisms Let $\mathcal{O}^{m \rightarrow k}$ be the set of mappings $\mathbf{g} : D^m \rightarrow D^k$. A mapping $\mathbf{g} \in \mathcal{O}^{m \rightarrow k}$ can also be viewed as a sequence of k operations $\mathbf{g} = (g_1, \dots, g_k)$ with $g_i \in \mathcal{O}^m$. We define a *generalized fractional polymorphism of arity $m \rightarrow k$* as a probability distribution ρ over $\mathcal{O}^{m \rightarrow k}$. We say that language Γ admits ρ if every cost function $f \in \Gamma$ of arity n satisfies

$$\sum_{\mathbf{g} \in \mathcal{O}^{m \rightarrow k}} \rho(\mathbf{g}) f^k(\mathbf{g}(x^1, \dots, x^m)) \leq f^m(x^1, \dots, x^m) \quad \forall x^1, \dots, x^m \in D^n \quad (2)$$

Equivalently, ρ is a generalized fractional polymorphism of Γ of arity $m \rightarrow k$ iff vector

$$\omega = \sum_{\mathbf{g}=(g_1, \dots, g_k) \in \mathcal{O}^{m \rightarrow k}} \rho(\mathbf{g}) \cdot \frac{1}{k} (\chi_{g_1} + \dots + \chi_{g_k}) \quad (3)$$

is a fractional polymorphism of Γ of arity m .

We can identify fractional polymorphisms of arity m with generalized fractional polymorphisms of arity $m \rightarrow 1$. For brevity, we will omit the word “generalized”.

We always use the following convention: if $\mathbf{g} = (g_1, \dots, g_k)$ is a mapping in $\mathcal{O}^{m \rightarrow k}$ then we extend it to m labelings $x^1, \dots, x^m \in D^n$ component-wise, i.e. $[\mathbf{g}(x^1, \dots, x^m)]_v = \mathbf{g}(x_v^1, \dots, x_v^m)$ for all $v \in [1, n]$. Thus, $\mathbf{g}(x^1, \dots, x^m)$ is a sequence of k labelings $(g_1(x^1, \dots, x^m), \dots, g_k(x^1, \dots, x^m))$ in D^n .

Fractional polymorphisms and SDP/LP relaxations It has been shown that fractional polymorphisms can be used for characterizing languages that can be solved exactly by certain convex relaxations. For the SDP relaxation, the following is known; it is implied by results in Chapters 6 and 7 of [14].

Theorem 2 ([14]). *If Γ has a cyclic fractional polymorphism of some arity $k \geq 2$ then the basic SDP relaxation solves $\text{VCSP}(\Gamma)$ in polynomial time.*

The more relevant for our paper case of the LP relaxation has been analyzed in [17]:^{1 2}

Theorem 3 ([17]). *For a finite-valued language Γ , the basic LP relaxation solves $\text{VCSP}(\Gamma)$ (in polynomial time) iff Γ admits an m -ary symmetric fractional polymorphism for every $m \geq 2$.*

Theorem 4 ([17], Theorem 4.4). *If Γ admits a k -ary fractional polymorphism ω such that $\text{supp}(\omega)$ generates a symmetric m -ary operation then Γ admits an m -ary symmetric fractional polymorphism.*

While the theorems give a necessary and sufficient condition for BLP to solve $\text{VCSP}(\Gamma)$, it was not clear whether checking this condition for a given language Γ is a decidable problem (we would need to consider infinitely many values of m).

Our result given below resolves this question affirmatively.

Theorem 5. *Suppose that a finite-valued language Γ admits a fractional polymorphism of arity $k \geq 2$ whose support contains at least one symmetric operation. Then Γ admits a symmetric fractional polymorphism of every arity $m \geq 2$ (and thus basic LP solves $\text{VCSP}(\Gamma)$).*

The question whether the existence of a cyclic fractional polymorphism implies the existence of a symmetric fractional polymorphism remains open.

¹Note, direction (i) \Rightarrow (ii) of theorem 3 was proved only for a language Γ with finite $|\Gamma|$. With a straightforward continuity argument the theorem can be extended to languages with countable Γ (by showing first that the set of m -ary fractional polymorphisms of any Γ is a compact subset of $\mathbb{R}^{\mathcal{O}^m}$).

²Thapper and Živný also present some results for infinite-valued languages; we refer to [17] for details.

STP multimorphisms As a minor contribution, we formally prove that if a finite-valued language admits an *STP multimorphism* then it also admits a submodularity multimorphism with respect to some total order on D . (This has implications for the complexity classification of *conservative* finite-valued languages.) This result is already known, but to our knowledge a formal proof has not been written down yet. We refer to section 4 for further discussion.

3 Proof of theorem 5

We will prove the following result.

Theorem 6. *Suppose that a finite-valued language Γ over a domain D admits a symmetric fractional polymorphism of arity $m - 1 \geq 2$. Then Γ admits a symmetric fractional polymorphism of arity m .*

This will imply theorem 5. Indeed, suppose that Γ admits a fractional polymorphism of arity $k \geq 2$ which contains a symmetric operation. Induction on m yields that Γ admits an m -ary symmetric fractional polymorphism for every $m \geq k$ (we need to use theorem 4 for the base case and theorem 6 for the induction step). This also implies the claim for all $m \in [2, k - 1]$: it is straightforward to show that if Γ admits a symmetric fractional polymorphism of arity pm where $p, m \in \mathbb{N}$ then it also admits a symmetric fractional polymorphism of arity m .

We thus concentrate on proving theorem 6. From now on we fix a symmetric fractional polymorphism ω of Γ of arity $m - 1$.

Consider a permutation π of size m and a symmetric operation $s \in \text{supp}(\omega)$ of arity $m - 1$. For such π and s we introduce the following definitions.

- For a labeling $\alpha = (a_1, \dots, a_m) \in D^m$, let $\alpha^\pi \in D^m$ and $\alpha^s \in D^m$ be the following labelings:

$$\alpha^\pi = (a_{\pi(1)}, \dots, a_{\pi(m)}) \quad (4)$$

$$\alpha^s = (s(\alpha_{-1}), \dots, s(\alpha_{-m})) \quad (5)$$

where $\alpha_{-i} \in D^{m-1}$ is the labeling obtained from α by removing the i -th element.

- For an operation $g : D^m \rightarrow D$, let $g^\pi : D^m \rightarrow D$ be the following operation:

$$g^\pi(\alpha) = g(\alpha^\pi) \quad (6)$$

- For a mapping $\mathbf{g} : D^m \rightarrow D^m$, let $\mathbf{g}^s : D^m \rightarrow D^m$ be the following mapping:

$$\mathbf{g}^s(\alpha) = [\mathbf{g}(\alpha)]^s \quad (7)$$

The last definition can also be expressed as

$$\mathbf{g}^s = (s \circ \mathbf{g}_{-1}, \dots, s \circ \mathbf{g}_{-m}) \quad (8)$$

where $\mathbf{g}_{-i} : D^m \rightarrow D^{m-1}$ is the sequence of $m - 1$ operations obtained from $\mathbf{g} = (g_1, \dots, g_m)$ by removing the i -th operation.

Let $\mathbb{1}$ be the identity mapping $D^m \rightarrow D^m$, and let $\mathcal{H}_\circ = \{\mathbb{1}^{s_1 \dots s_k} \mid s_1, \dots, s_k \in \text{supp}(\omega), k \geq 0\}$ be the set of all mappings that can be generated from $\mathbb{1}$.

Proposition 7. *Every $\mathbf{g} = (g_1, \dots, g_m) \in \mathcal{H}_\circ$ satisfies the following:*

$$(g_1^\pi, \dots, g_m^\pi) = (g_{\pi(1)}, \dots, g_{\pi(m)}) \quad \forall \text{ permutation } \pi \quad (9)$$

Thus, permuting the arguments of $g_i(\cdot, \dots, \cdot)$ gives a mapping which is also present in the sequence \mathbf{g} , possibly at a different position.

Proof. Checking that $\mathbb{1}$ satisfies (9) is straightforward. Let us prove that for any $\mathbf{g} : D^m \rightarrow D^m$ satisfying (9) and for any symmetric operation $s \in \mathcal{O}^{m-1}$ mapping \mathbf{g}^s also satisfies (9). Consider $i \in [1, m]$. We need to show that $(s \circ \mathbf{g}_{-i})^\pi = s \circ \mathbf{g}_{-\pi(i)}$. For each $\alpha \in D^m$ we have

$$\begin{aligned} (s \circ \mathbf{g}_{-i})^\pi(\alpha) &= s \circ \mathbf{g}_{-i}(\alpha^\pi) = s(g_1(\alpha^\pi), \dots, g_{i-1}(\alpha^\pi), g_{i+1}(\alpha^\pi), \dots, g_m(\alpha^\pi)) \\ &= s(g_1^\pi(\alpha), \dots, g_{i-1}^\pi(\alpha), g_{i+1}^\pi(\alpha), \dots, g_m^\pi(\alpha)) \\ &= s(g_{\pi(1)}(\alpha), \dots, g_{\pi(i-1)}(\alpha), g_{\pi(i+1)}(\alpha), \dots, g_{\pi(m)}(\alpha)) = s \circ \mathbf{g}_{-\pi(i)}(\alpha) \end{aligned}$$

□

Graph on mappings Let us define a directed weighted graph $G = (\mathcal{H}_\circ, E, w)$ with the set of edges $E = \{(\mathbf{g}, \mathbf{g}^s) \mid \mathbf{g} \in \mathcal{H}_\circ, s \in \text{supp}(\omega)\}$ and positive weights $w(\mathbf{g}, \mathbf{h}) = \sum_{s \in \text{supp}(\omega), \mathbf{h} = \mathbf{g}^s} \omega(s)$ for $(\mathbf{g}, \mathbf{h}) \in E$. Clearly, we have

$$\sum_{(\mathbf{g}, \mathbf{h}) \in E} w(\mathbf{g}, \mathbf{h}) = 1 \quad \mathbf{g} \in \mathcal{H}_\circ \quad (10)$$

We define $\mathbb{H}[G]$ to be the set of strongly connected components $\mathcal{H} \subseteq \mathcal{H}_\circ$ of G which are sinks, i.e. all edges in G from \mathcal{H} lead to vertices in \mathcal{H} .

Proposition 8. *For any $\mathcal{H} \in \mathbb{H}[G]$ and $\hat{\mathbf{g}} \in \mathcal{H}$ there exists $\mathbf{g} \in \mathcal{H}$ with $\mathbf{g} \circ \hat{\mathbf{g}} = \hat{\mathbf{g}}$.*

Proof. Note that $\mathbb{1}^{s_1 \dots s_k} \circ \mathbf{h} = \mathbf{h}^{s_1 \dots s_k}$ for any $s_1, \dots, s_k \in \text{supp}(\omega)$ and $\mathbf{h} \in \mathcal{O}^{m \rightarrow m}$, since

$$[\mathbb{1}^{s_1 \dots s_k} \circ \mathbf{h}](\alpha) = \mathbb{1}^{s_1 \dots s_k}(\mathbf{h}(\alpha)) = [\mathbb{1}(\mathbf{h}(\alpha))]^{s_1 \dots s_k} = [\mathbf{h}(\alpha)]^{s_1 \dots s_k} = \mathbf{h}^{s_1 \dots s_k}(\alpha) \quad \forall \alpha \in D^m$$

Therefore, conditions $\mathbf{g} \in \mathcal{H}_\circ$, $\mathbf{h} \in \mathcal{H}$ imply that $\mathbf{g} \circ \mathbf{h} \in \mathcal{H}$ (since \mathbf{g} can be written as $\mathbf{g} = \mathbb{1}^{s_1 \dots s_k}$ and there are no edges leaving \mathcal{H}).

Since \mathcal{H} is strongly connected, there is a path in G from $\hat{\mathbf{g}} \circ \hat{\mathbf{g}} \in \mathcal{H}$ to $\hat{\mathbf{g}} \in \mathcal{H}$, i.e. $[\hat{\mathbf{g}} \circ \hat{\mathbf{g}}]^{s_1 \dots s_k} = \hat{\mathbf{g}}$ for some $s_1, \dots, s_k \in \text{supp}(\omega)$. Equivalently, $\mathbf{h} \circ \hat{\mathbf{g}} \circ \hat{\mathbf{g}} = \hat{\mathbf{g}}$ where $\mathbf{h} = \mathbb{1}^{s_1 \dots s_k}$. It can be checked that mapping $\mathbf{g} = \mathbf{h} \circ \hat{\mathbf{g}}$ has the desired properties. □

3.1 Proof overview: main theorems

From now on we fix a function $f \in \Gamma$ of arity n . Recall that we defined $f^m(x^1, \dots, x^m) = \frac{1}{m} \sum_{i \in [1, m]} f(x^i)$. First, we will prove

Theorem 9. *For each $\mathcal{H} \in \mathbb{H}[G]$ there exists a probability distribution λ over \mathcal{H} (i.e. vector $\lambda \in \mathbb{R}_{\geq 0}^{\mathcal{H}}$ with $\sum_{\mathbf{g} \in \mathcal{H}} \lambda_{\mathbf{g}} = 1$) that satisfies the following for any labelings $x^1, \dots, x^m \in D^n$.*

Denote $I = \{(\mathbf{g}, i) \mid \mathbf{g} \in \mathcal{H}, i \in [1, m]\}$, and for each $(\mathbf{g}, i) \in I$ define labeling $x^{\mathbf{g}^i} \in D^n$ via

$$(x^{\mathbf{g}^1}, \dots, x^{\mathbf{g}^m}) = (\mathbf{g}(x^1), \dots, \mathbf{g}(x^m)) \quad (11)$$

(a) *For any $k \in [1, m]$ there holds*

$$\sum_{i \in [1, m]} c_i^k \left[\sum_{\mathbf{g} \in \mathcal{H}} \lambda_{\mathbf{g}} \cdot f(x^{\mathbf{g}^i}) \right] \leq 0 \quad \text{where} \quad c_i^k = \begin{cases} 1 & \text{if } i = k \\ -\frac{1}{m-1} & \text{if } i \in [1, m] - \{k\} \end{cases} \quad (12)$$

(b) *If $\mathbf{g}', \mathbf{g}'' \in \mathcal{H}$ then $f^m(x^{\mathbf{g}'^1}, \dots, x^{\mathbf{g}'^m}) = f^m(x^{\mathbf{g}''^1}, \dots, x^{\mathbf{g}''^m})$.*

Using this technical result, we will then establish

Theorem 10. *Consider a strongly connected component $\mathcal{H} \in \mathbb{H}[G]$ and a mapping $\mathbf{g} \in \mathcal{H}$. Denote $\text{Range}_n(\mathbf{g}) = \{\mathbf{g}(x^1, \dots, x^m) \mid x^1, \dots, x^m \in D^n\}$, and let $\mathbf{p} \in \mathcal{O}^{m \rightarrow m}$ be any mapping such that $\mathbf{p}(\alpha)$ is a permutation of α for all $\alpha \in D^m$.*

For any $(x^1, \dots, x^m) \in \text{Range}_n(\mathbf{g})$ there holds $f^m(x^1, \dots, x^m) = f^m(\mathbf{p}(x^1, \dots, x^m))$.

Finally, we will prove the following.

Theorem 11. *Denote $\overline{H} = \bigcup_{\mathcal{H} \in \mathbb{H}[G]} \subseteq \mathcal{H}_\circ$. There exists a fractional polymorphism ρ of Γ of arity $m \rightarrow m$ with $\text{supp}(\rho) \subseteq \overline{H}$.*

Proving theorems 10 and 11 will be sufficient for establishing theorem 6. Indeed, we can construct an m -ary symmetric fractional polymorphism of Γ as follows. Take a symmetric mapping $\mathbf{p} \in \mathcal{O}^{m \rightarrow m}$ satisfying the condition of theorem 10, take vector ρ from theorem 11, and define a fractional polymorphism of arity $m \rightarrow m$

$$\rho' = \sum_{\mathbf{g} \in \text{supp}(\rho)} \rho(\mathbf{g}) \chi_{\mathbf{p} \circ \mathbf{g}}$$

Language Γ admits ρ' since for any labelings $x^1, \dots, x^m \in D^m$ there holds

$$\begin{aligned} \sum_{\mathbf{h} \in \text{supp}(\rho')} \rho'(\mathbf{h}) f^m(\mathbf{h}(x^1, \dots, x^m)) &= \sum_{\mathbf{g} \in \text{supp}(\rho)} \rho(\mathbf{g}) f^m(\mathbf{p}(\mathbf{g}(x^1, \dots, x^m))) \\ &= \sum_{\mathbf{g} \in \text{supp}(\rho)} \rho(\mathbf{g}) f^m(\mathbf{g}(x^1, \dots, x^m)) \leq f^m(x^1, \dots, x^m) \end{aligned}$$

Note, for any $\mathbf{h} = (h_1, \dots, h_m) \in \text{supp}(\rho')$ operations h_1, \dots, h_m are symmetric. Indeed, we have $\mathbf{h} = \mathbf{p} \circ \mathbf{g}$ for some $\mathbf{g} \in \mathcal{H}_\circ$. If $\alpha \in D^m$ and π is a permutation of size m then $\mathbf{h}(\alpha^\pi) = \mathbf{p}(\mathbf{g}(\alpha^\pi)) = \mathbf{p}([\mathbf{g}(\alpha)]^\pi) = \mathbf{p}(\mathbf{g}(\alpha)) = \mathbf{h}(\alpha)$ which implies the claim.

We can finally apply transformation (3) to vector ρ' to get a symmetric fractional polymorphism of arity m .

3.2 Proof of theorem 9

First, we make the following observation.

Proposition 12. *If $\mathbf{h} = \mathbf{g}^s$ where $\mathbf{g} \in \mathcal{H}$, $s \in \text{supp}(\omega)$ then $x^{\mathbf{h}i} = s((x^{\mathbf{g}^1}, \dots, x^{\mathbf{g}^m})_{-i})$ for $i \in [1, m]$ where $(x^{\mathbf{g}^1}, \dots, x^{\mathbf{g}^m})_{-i}$ is the sequence of $m-1$ labelings obtained by removing the i -th labeling.*

Proof. Consider node $v \in [1, n]$, and denote $\alpha = (x_v^1, \dots, x_v^m)$, $\beta = (x_v^{\mathbf{g}^1}, \dots, x_v^{\mathbf{g}^m})$, $\gamma = (x_v^{\mathbf{h}^1}, \dots, x_v^{\mathbf{h}^m})$. By definition (11), $\beta = \mathbf{g}(\alpha)$ and $\gamma = \mathbf{h}(\alpha)$. Therefore, $\gamma = \mathbf{g}^s(\alpha) = [\mathbf{g}(\alpha)]^s = \beta^s$. In other words, the i -th component of γ equals $s(\beta_{-i})$, which is what we needed to show. \square

For each $(\mathbf{g}, i) \in I$ we can write the polymorphism inequality for labelings $(x^{\mathbf{g}^1}, \dots, x^{\mathbf{g}^m})_{-i}$:

$$\sum_{s \in \text{supp}(\omega)} \omega(s) f(s((x^{\mathbf{g}^1}, \dots, x^{\mathbf{g}^m})_{-i})) \leq \frac{1}{m-1} \sum_{j \in [1, m] - \{i\}} f(x^{\mathbf{g}^j}) \quad (13)$$

Let us multiply this inequality by weight $\lambda_{\mathbf{g}i} \geq 0$ (to be defined later), and apply proposition 12:

$$\lambda_{\mathbf{g}i} \sum_{s \in \text{supp}(\omega), \mathbf{h} = \mathbf{g}^s} \omega(s) f(x^{\mathbf{h}i}) - \frac{\lambda_{\mathbf{g}i}}{m-1} \sum_{j \in [1, m] - \{i\}} f(x^{\mathbf{g}^j}) \leq 0 \quad \forall (\mathbf{g}, i) \in I \quad (14)$$

Summing these inequalities over $(\mathbf{g}, i) \in I$ gives

$$\sum_{(\mathbf{g}, i) \in I} \left[\sum_{s \in \text{supp}(\omega), \mathbf{g}=\mathbf{h}^s} \omega(s) \lambda_{\mathbf{h}i} - \sum_{j \in [1, m] - \{i\}} \frac{\lambda_{\mathbf{g}j}}{m-1} \right] f(x^{\mathbf{g}i}) \leq 0 \quad \forall (\mathbf{g}, i) \in I \quad (15)$$

The last inequality is equivalent to (12) assuming that the following holds:

$$\sum_{(\mathbf{h}, \mathbf{g}) \in E} w(\mathbf{h}, \mathbf{g}) \lambda_{\mathbf{h}i} - \sum_{j \in [1, m] - \{i\}} \frac{\lambda_{\mathbf{g}j}}{m-1} = c_i \lambda_{\mathbf{g}} \quad \forall (\mathbf{g}, i) \in I \quad (16)$$

where we denoted $c_i = c_i^k$ for brevity. Next, we show that eq. (16) can be satisfied with the appropriate choice of weights $\{\lambda_{\mathbf{g}i}, \lambda_{\mathbf{g}}\}$, thus proving theorem 9(a).

Lemma 13. *There exists vector $\lambda \in \mathbb{R}_{\geq 0}^{I \cup \mathcal{H}}$ that satisfies (16) and*

$$\sum_{\mathbf{g} \in \mathcal{H}} \lambda_{\mathbf{g}} = 1 \quad (17)$$

Proof. Suppose the lemma does not hold. By Farkas lemma there exist vector $y \in \mathbb{R}^I$ and scalar $z \in \mathbb{R}$ such that

$$z < 0 \quad (18a)$$

$$z - \sum_{i \in [1, m]} c_i y_{\mathbf{g}i} \geq 0 \quad \forall \mathbf{g} \in \mathcal{H} \quad (18b)$$

$$\sum_{(\mathbf{g}, \mathbf{h}) \in E} w(\mathbf{g}, \mathbf{h}) y_{\mathbf{h}i} - \sum_{j \in [1, m] - \{i\}} \frac{y_{\mathbf{g}j}}{m-1} \geq 0 \quad \forall (\mathbf{g}, i) \in I \quad (18c)$$

Denote $u_{\mathbf{g}} = \sum_{i \in [1, m]} y_{\mathbf{g}i}$. Summing inequalities (18c) over $i \in [1, m]$ gives

$$\sum_{(\mathbf{g}, \mathbf{h}) \in E} w(\mathbf{g}, \mathbf{h}) u_{\mathbf{h}} - u_{\mathbf{g}} \geq 0 \quad \forall \mathbf{g} \in \mathcal{H} \quad (19)$$

Denote $\mathcal{H}^* = \arg \max \{u_{\mathbf{g}} \mid \mathbf{g} \in \mathcal{H}\}$. From (10) and (19) we conclude that $\mathbf{g} \in \mathcal{H}^*$ implies $\mathbf{h} \in \mathcal{H}^*$ for all $(\mathbf{g}, \mathbf{h}) \in E$. Therefore, $\mathcal{H}^* = \mathcal{H}$ (since \mathcal{H} is a strongly connected component of G).

We showed that $u_{\mathbf{g}} = C$ for all $\mathbf{g} \in \mathcal{H}$ where $C \in \mathbb{R}$ is some constant. We can assume w.l.o.g. that this constant is zero. Indeed, this can be achieved by subtracting C/m from values $y_{\mathbf{g}i}$ for all $(\mathbf{g}, i) \in I$ with $\mathbf{g} \in \mathcal{H}$; it can be checked (using eq. (10)) that this operation preserves inequalities (18). We thus have

$$\sum_{j \in [1, m] - \{i\}} y_{\mathbf{g}j} = -y_{\mathbf{g}i} \quad \forall (\mathbf{g}, i) \in I \quad (20)$$

Substituting this into (18c) gives

$$\sum_{(\mathbf{g}, \mathbf{h}) \in E} w(\mathbf{g}, \mathbf{h}) y_{\mathbf{h}i} + \frac{y_{\mathbf{g}i}}{m-1} \geq 0 \quad \forall (\mathbf{g}, i) \in I \quad (21)$$

Consider $i' \in [1, m]$. Summing (21) over $i \in [1, m] - \{i'\}$ and then using (20) yields

$$\sum_{(\mathbf{g}, \mathbf{h}) \in E} w(\mathbf{g}, \mathbf{h}) (-y_{\mathbf{h}i'}) + \frac{-y_{\mathbf{g}i'}}{m-1} \geq 0 \quad \forall (\mathbf{g}, i') \in I \quad (22)$$

Combining (21) and (22) gives

$$\sum_{(\mathbf{g}, \mathbf{h}) \in E} w(\mathbf{g}, \mathbf{h}) y_{\mathbf{h}i} + \frac{y_{\mathbf{g}i}}{m-1} = 0 \quad \forall (\mathbf{g}, i) \in I \quad (23)$$

Denote $r_{\mathbf{g}} = \sum_{i \in [1, m]} c_i y_{\mathbf{g}i}$ for $\mathbf{g} \in \mathcal{H}$. Summing (23) over $i \in [1, m]$ with appropriate coefficients gives

$$\sum_{(\mathbf{g}, \mathbf{h}) \in E} w(\mathbf{g}, \mathbf{h}) r_{\mathbf{h}} + \frac{r_{\mathbf{g}}}{m-1} = 0 \quad \forall \mathbf{g} \in \mathcal{H} \quad (24)$$

From (18a) and (18b) we conclude that $r_{\mathbf{g}} < 0$ for all $\mathbf{g} \in \mathcal{H}$, and thus eq. (24) cannot hold - a contradiction. \square

We have proved part (a) of theorem 9. Next, we use a similar strategy to establish part (b). Namely, lemma 14 below shows that we can choose weights $\{\lambda_{\mathbf{g}i}\}$ so that inequality (15) becomes equivalent to

$$\sum_{i \in [1, m]} f(x^{\mathbf{g}'i}) - \sum_{i \in [1, m]} f(x^{\mathbf{g}''i}) \leq 0 \quad (25)$$

Using these inequalities for pairs $(\mathbf{g}', \mathbf{g}'')$ and $(\mathbf{g}'', \mathbf{g}')$ then yields theorem 9(b).

Lemma 14. *There exists vector $\lambda \in \mathbb{R}_{\geq 0}^I$ that satisfies*

$$\sum_{(\mathbf{h}, \mathbf{g}) \in E} w(\mathbf{h}, \mathbf{g}) \lambda_{\mathbf{h}i} - \sum_{j \in [1, m] - \{i\}} \frac{\lambda_{\mathbf{g}j}}{m-1} = c_{\mathbf{g}i} \quad \forall (\mathbf{g}, i) \in I \quad (26)$$

where vector $c \in \mathbb{R}^I$ is defined via

$$c_{\mathbf{g}i} = \begin{cases} 1 & \text{if } \mathbf{g} = \mathbf{g}' \\ -1 & \text{if } \mathbf{g} = \mathbf{g}'' \\ 0 & \text{otherwise} \end{cases} \quad (27)$$

Proof. Suppose the lemma does not hold. By Farkas lemma there exist vector $y \in \mathbb{R}^I$ such that

$$\sum_{(\mathbf{g}, i) \in I} c_{\mathbf{g}i} y_{\mathbf{g}i} < 0 \quad (28a)$$

$$\sum_{(\mathbf{g}, \mathbf{h}) \in E} w(\mathbf{g}, \mathbf{h}) y_{\mathbf{h}i} - \sum_{j \in [1, m] - \{i\}} \frac{y_{\mathbf{g}j}}{m-1} \geq 0 \quad \forall (\mathbf{g}, i) \in I \quad (28b)$$

Denote $u_{\mathbf{g}} = \sum_{i \in [1, m]} y_{\mathbf{g}i}$. Using the same argument as in the proof of lemma 13 we conclude that $u_{\mathbf{g}} = C$ for all $\mathbf{g} \in \mathcal{H}$ where $C \in \mathbb{R}$ is some constant. But then the expression on the LHS of (28a) equals $C - C = 0$ - a contradiction. \square

3.3 Proof of theorem 10

Let $\lambda \in \mathbb{R}_{\geq 0}^{\mathcal{H}}$ be a vector constructed in theorem 9. Define

$$F_i(x^1, \dots, x^m) = \sum_{\mathbf{g} \in \mathcal{H}} \lambda_{\mathbf{g}} \cdot f(x^{\mathbf{g}i}) \quad \text{where } (x^{\mathbf{g}1}, \dots, x^{\mathbf{g}m}) = \mathbf{g}(x^1, \dots, x^m) \quad (29)$$

By theorem 9(a) we have

$$\sum_{i \in [1, m]} c_i^k \cdot F_i(x^1, \dots, x^m) \leq 0 \quad \forall k \in [1, m] \quad (30)$$

Summing these inequalities over $k \in [1, m]$ gives 0 on the LHS. Therefore, each inequality in (30) is actually an equality. Subtracting these equalities for indices k and k' gives

$$F_{k'}(x^1, \dots, x^m) = F_k(x^1, \dots, x^m) \quad \forall k, k' \in [1, m] \quad (31)$$

We denote $F(x^1, \dots, x^m) = F_i(x^1, \dots, x^m)$ for $i \in [1, m]$.

Proposition 15. *The following transformation does not change $F(x^1, \dots, x^m)$: pick node $v \in [1, m]$ and permute values (x_v^1, \dots, x_v^m) .*

Proof. It suffices to prove the claim for a permutation which swaps values x_v^i and x_v^j for $i, j \in [1, m]$ (since any other permutation can be obtained by repeatedly applying such swaps). Since $m \geq 3$ there exists index $k \in [1, m] - \{i, j\}$. Using proposition 7, it can be checked that the swap above does not affect labelings $x^{\mathbf{g}^k}$ in (29) for $\mathbf{g} \in \mathcal{H}$, and therefore $F_k(x^1, \dots, x^m)$ does not change. \square

Now suppose that $(x^1, \dots, x^m) \in \text{Range}_n(\hat{\mathbf{g}})$ where $\hat{\mathbf{g}} \in \mathcal{H}$. By proposition 8 there exists $\mathbf{g} \in \mathcal{H}$ with $\mathbf{g} \circ \hat{\mathbf{g}} = \hat{\mathbf{g}}$. Clearly, we have $\mathbf{g}(x^1, \dots, x^m) = (x^1, \dots, x^m)$, or equivalently $(x^{\mathbf{g}^1}, \dots, x^{\mathbf{g}^m}) = (x^1, \dots, x^m)$. Therefore, theorem 9(b) implies that $f^m(x^{\mathbf{g}^1}, \dots, x^{\mathbf{g}^m}) = f^m(x^1, \dots, x^m)$ for all $\mathbf{g} \in \mathcal{H}$. Using this fact and the definition of $F_i(\cdot)$, we can write

$$\begin{aligned} F(x^1, \dots, x^m) &= \frac{1}{m} \sum_{i \in [1, m]} F_i(x^1, \dots, x^m) = \frac{1}{m} \sum_{\mathbf{g} \in \mathcal{H}} \lambda_{\mathbf{g}} \sum_{i \in [1, m]} f(x^{\mathbf{g}^i}) \\ &= \sum_{\mathbf{g} \in \mathcal{H}} \lambda_{\mathbf{g}} f^m(x^{\mathbf{g}^1}, \dots, x^{\mathbf{g}^m}) = \sum_{\mathbf{g} \in \mathcal{H}} \lambda_{\mathbf{g}} f^m(x^1, \dots, x^m) = f^m(x^1, \dots, x^m) \end{aligned}$$

To establish theorem 10, it remains to prove that condition $(x^1, \dots, x^m) \in \text{Range}_n(\hat{\mathbf{g}})$ implies $\mathbf{p}(x^1, \dots, x^m) \in \text{Range}_n(\hat{\mathbf{g}})$. This proof follows mechanically from proposition 7, and is omitted.

3.4 Proof of theorem 11

We will rely on the following fact.

Proposition 16. *Suppose that ρ is a fractional polymorphism of Γ of arity $m \rightarrow m$ and $\mathbf{g} \in \text{supp}(\rho)$ with $\lambda = \rho(\mathbf{g})$. Then vector*

$$\rho' = \rho - \lambda \cdot \chi_{\mathbf{g}} + \lambda \sum_{s \in \text{supp}(\omega)} \omega(s) \chi_{\mathbf{g}^s} \quad (32)$$

is also a fractional polymorphism of Γ of arity $m \rightarrow m$.

Proof. Denote the sum on the LHS of (32) as $\rho_{\mathbf{g}}$, so that $\rho' = \rho - \lambda \cdot \chi_{\mathbf{g}} + \lambda \cdot \rho_{\mathbf{g}}$. Consider function $f \in \Gamma$ of arity n and labelings $x^1, \dots, x^m \in D^n$, and denote $(y^1, \dots, y^m) = \mathbf{g}(x^1, \dots, x^m)$. We can write

$$\begin{aligned} \sum_{\mathbf{h} \in \text{supp}(\rho_{\mathbf{g}})} \rho_{\mathbf{g}}(\mathbf{h}) f^m(\mathbf{h}(x^1, \dots, x^m)) &= \sum_{s \in \text{supp}(\omega)} \omega(s) f^m(\mathbf{g}^s(x^1, \dots, x^m)) \\ &= \sum_{s \in \text{supp}(\omega)} \omega(s) \frac{1}{m} \sum_{i \in [1, m]} f((y^1, \dots, y^m)_{-i}) \\ &= \frac{1}{m} \sum_{i \in [1, m]} \sum_{s \in \text{supp}(\omega)} \omega(s) f((y^1, \dots, y^m)_{-i}) \\ &\leq \frac{1}{m} \sum_{i \in [1, m]} f^{m-1}((y^1, \dots, y^m)_{-i}) \\ &= f^m(y^1, \dots, y^m) = f^m(\mathbf{g}(x^1, \dots, x^m)) \end{aligned}$$

Therefore,

$$\begin{aligned}
\sum_{\mathbf{h} \in \text{supp}(\rho')} \rho'(\mathbf{h}) f^m(\mathbf{h}(x^1, \dots, x^m)) &= \sum_{\mathbf{h} \in \text{supp}(\rho) - \{\mathbf{g}\}} \rho(\mathbf{h}) f^m(\mathbf{h}(x^1, \dots, x^m)) + \lambda \sum_{\mathbf{h} \in \text{supp}(\rho_{\mathbf{g}})} \rho_{\mathbf{g}}(\mathbf{h}) f^m(\mathbf{h}(x^1, \dots, x^m)) \\
&\leq \sum_{\mathbf{h} \in \text{supp}(\rho) - \{\mathbf{g}\}} \rho(\mathbf{h}) f^m(\mathbf{h}(x^1, \dots, x^m)) + \rho(\mathbf{g}) f^m(\mathbf{g}(x^1, \dots, x^m)) \\
&= \sum_{\mathbf{h} \in \text{supp}(\rho)} \rho(\mathbf{h}) f^m(\mathbf{h}(x^1, \dots, x^m)) \leq f^m(x^1, \dots, x^m)
\end{aligned}$$

□

We can now prove theorem 11. Let Ω be the set of fractional polymorphisms ρ of Γ of arity $m \rightarrow m$ with $\text{supp}(\rho) \subseteq \mathcal{H}_\circ$. Let us pick $\rho \in \Omega$ with the maximum value of $\rho(\overline{\mathcal{H}}) \doteq \sum_{\mathbf{g} \in \overline{\mathcal{H}}} \rho(\mathbf{g})$. (Clearly, Ω is a compact set, so the maximum is attained by some vector in Ω). We claim that $\text{supp}(\rho) \in \overline{\mathcal{H}}$. Indeed, suppose that there exists $\mathbf{g} \in \text{supp}(\rho)$ with $\mathbf{g} \notin \overline{\mathcal{H}}$. By the definition of $\overline{\mathcal{H}}$ there exists a path in G from \mathbf{g} to a vertex $\mathbf{h} \in \overline{\mathcal{H}}$, i.e. $\mathbf{h} = \mathbf{g}^{s_1 \dots s_k}$ for some $s_1, \dots, s_k \in \text{supp}(\omega)$. Therefore, we can repeatedly modify ρ by applying a sequence of transformations (32) to the current vector ρ and some $\mathbf{g}_i \in \text{supp}(\rho) - \overline{\mathcal{H}}$ to get vector ρ' with larger weight $\rho'(\overline{\mathcal{H}})$, using at most k steps. This contradicts to the choice of ρ . Theorem 11 is proved.

We mention that a similar argument can be used as an alternative to the “tree cutting” argument in the proof of theorem 4.4 in [17] (and vice versa).

4 STP multimorphisms

In this section we consider Symmetric Tournament Pairs (STP) multimorphisms [3].

Definition 17. (a) A pair of operations $\langle \sqcap, \sqcup \rangle$ with $\sqcap, \sqcup : D \times D \rightarrow D$ is called an STP multimorphism if

$$a \sqcap a = a \sqcup a = a \quad \forall a \in D \quad (\text{idempotency}) \quad (33a)$$

$$a \sqcap b = b \sqcap a, \quad a \sqcup b = b \sqcup a \quad \forall a, b \in D \quad (\text{commutativity}) \quad (33b)$$

$$\{a \sqcap b, a \sqcup b\} = \{a, b\} \quad \forall a, b \in D \quad (\text{conservativity}) \quad (33c)$$

(b) Pair $\langle \sqcap, \sqcup \rangle$ is called a submodularity multimorphism if there exists a total order on D for which $a \sqcap b = \min\{a, b\}$, $a \sqcup b = \max\{a, b\}$ for all $a, b \in D$.

(c) Language Γ admits $\langle \sqcap, \sqcup \rangle$ if every function $f \in \Gamma$ of arity n satisfies

$$f(x \sqcap y) + f(x \sqcup y) \leq f(x) + f(y) \quad \forall x, y \in D^n \quad (34)$$

It has been shown in [3] that if Γ admits an STP multimorphism then $VCSP(\Gamma)$ can be solved in polynomial time. STP multimorphisms also appeared in the dichotomy result of [10]:

Theorem 18. Suppose a finite-valued language Γ is conservative, i.e. it contains all possible unary cost functions $u : D \rightarrow \{0, 1\}$. Then Γ either admits an STP multimorphism or it is NP-hard.

To goal of this section is to prove the following.

Theorem 19. If a finite-valued language Γ admits an STP multimorphism then it also admits a submodularity multimorphism.

This fact is already known; in particular, footnote 2 in [10] mentions that this result is implicitly contained in [3], and sketches a proof strategy. However, to our knowledge a formal proof has never appeared in the literature. We now fill this gap. Our proof is different from the one suggested in [10], and inspired some of the proof techniques used in the main part of this paper.

4.1 Proof of theorem 19

Consider a directed graph $G = (D, E)$. We say that G is *complete* if for each pair of distinct labels $a, b \in D$ exactly one of the edges (a, b) , (b, a) belong to E . We define a one-to-one correspondence between STP multimorphisms $\langle \sqcap, \sqcup \rangle$ and complete graphs $G = (D, E)$ as follows:

$$(a, b) \in E \quad \Leftrightarrow \quad (a \sqcap b, a \sqcup b) = (a, b) \quad \forall a, b \in D, a \neq b$$

It can be seen that $\langle \sqcap, \sqcup \rangle$ is a submodularity multimorphism iff the corresponding graph G is acyclic.

Lemma 20. *Suppose a finite-valued language Γ admits an STP multimorphism $\langle \sqcap, \sqcup \rangle$ corresponding to a directed graph $G = (D, E)$, and suppose that G has a 3-cycle: $(a, b), (b, c), (c, a) \in E$. Let \hat{G} be the graph obtained from G by reversing the orientation of edge (a, b) , and let $\langle \hat{\sqcap}, \hat{\sqcup} \rangle$ be the corresponding STP multimorphism. Then Γ admits $\langle \hat{\sqcap}, \hat{\sqcup} \rangle$.*

Proof. Let $\langle \wedge, \vee \rangle$ be the following multimorphism:

$$(x \wedge y, x \vee y) = \begin{cases} (x, y) & \text{if } (x, y) \in \{(a, b), (b, a)\} \\ (x \sqcap y, x \sqcup y) & \text{if } (x, y) \notin \{(a, b), (b, a)\} \end{cases}$$

First, we will prove that Γ admits $\langle \wedge, \vee \rangle$ (step 1), and then prove that Γ admits $\langle \hat{\sqcap}, \hat{\sqcup} \rangle$ (step 2). We fix below function $f \in \Gamma$ of arity n and labelings $x, y \in D^n$.

Step 1 Let us define labelings $x', y' \in D^n$ via

$$(x'_v, y'_v) = \begin{cases} (x_v, x_v \sqcap y_v) & \text{if } (x_v, y_v) \neq (b, a) \\ (c, c) & \text{if } (x_v, y_v) = (b, a) \end{cases} \quad \forall v \in [1, n]$$

It can be checked that the following identities hold:

$$x' \sqcap y = y' \quad x \sqcup y' = x' \quad (35a)$$

$$x \sqcap y' = x \wedge y \quad x' \sqcup y = x \vee y \quad (35b)$$

Let us write multimorphism inequalities for pairs (x', y) and (x, y') :

$$\underline{f(x' \sqcap y)} + f(x' \sqcup y) \leq \underline{f(x')} + f(y) \quad (36a)$$

$$f(x \sqcap y') + \underline{f(x \sqcup y')} \leq f(x) + \underline{f(y')} \quad (36b)$$

Summing (36a) and (36b), cancelling terms using (35a), and then substituting expressions using (35b) gives

$$f(x \wedge y) + f(x \vee y) \leq f(x) + f(y) \quad (37)$$

Step 2 Let us define labelings $x', y' \in D^n$ via

$$(x'_v, y'_v) = \begin{cases} (x_v \wedge y_v, y_v) & \text{if } (x_v, y_v) \neq (a, b) \\ (c, c) & \text{if } (x_v, y_v) = (a, b) \end{cases} \quad \forall v \in [1, n]$$

It can be checked that the following identities hold:

$$x' \vee y = y' \quad x \wedge y' = x' \quad (38a)$$

$$x \vee y' = x \hat{\sqcup} y \quad x' \wedge y = x \hat{\sqcap} y \quad (38b)$$

Let us write multimorphism inequalities for pairs (x', y) and (x, y') :

$$f(x' \wedge y) + \underline{f(x' \vee y)} \leq \underline{f(x')} + f(y) \quad (39a)$$

$$\underline{f(x \wedge y')} + f(x \vee y') \leq f(x) + \underline{\underline{f(y')}} \quad (39b)$$

Summing (39a) and (39b), cancelling terms using (38a), and then substituting expressions using (38b) gives

$$f(x \hat{\cap} y) + f(x \hat{\cup} y) \leq f(x) + f(y) \quad (40)$$

□

We call the operation of reversing the orientation of edge $(a, b) \in E$ in a graph $G = (D, E)$ a *valid flip* if (a, b) belongs to a 3-cycle. To prove theorem 19, it thus suffices to show the following:

- For any complete graph G there exists a sequence of valid flips that makes it acyclic.

Such sequence can be constructed as follows: (1) start with a subset $B \subseteq D$ with $|B| = 2$; (2) perform valid flips in $G[B]$ to make it acyclic, where $G[B] = (B, E[B])$ is the subgraph of G induced by B ; (3) add a node $c \in D - B$ to B (if it exists) and repeat step 2. The lemma below shows how to implement step 2.

Lemma 21. *Suppose that $G = (D, E)$ is a complete graph, $D = B \cup \{c\}$ with $c \notin B$ and subgraph $G[B]$ is acyclic. Then there exists a sequence of valid flips that makes G acyclic.*

Proof. Suppose that G has a cycle \mathcal{C} , then it must pass through c (since $G[B]$ is acyclic): $\mathcal{C} = \dots \rightarrow b \rightarrow c \rightarrow a \rightarrow \dots$. Since there is a path from a to b in $G[B]$, we must have $(a, b) \in E$ (again, due to acyclicity of $G[B]$). Thus, $c \rightarrow a \rightarrow b \rightarrow c$ is a 3-cycle in G .

Let us repeat the following procedure while possible: pick such cycle and flip edge (c, a) to (a, c) . This operation decreases the number of edges in G coming out of c . Therefore, it must terminate after a finite number of steps and yield an acyclic graph G . □

Acknowledgements

I thank Andrei Krokhin for helpful discussions and for communicating the result of Raghavendra [14] about cyclic fractional polymorphisms.

References

- [1] A. Blake, P. Kohli, and C. Rother, editors. *Advances in Markov Random Fields for Vision and Image Processing*. MIT Press, 2011.
- [2] D. Cohen, M. Cooper, and P. Jeavons. An algebraic characterisation of complexity for valued constraints. In *In: Proceedings CP06. Volume 4204 of Lecture Notes in Computer Science., Springer-Verlag*, pages 107–121, 2006.
- [3] D. Cohen, M. Cooper, and P. Jeavons. Generalising submodularity and Horn clauses: Tractable optimization problems defined by tournament pair multimorphisms. *Theoretical Computer Science*, 401(1):36–51, 2008.
- [4] D. Cohen, M. Cooper, P. Jeavons, and A. Krokhin. The complexity of soft constraint satisfaction. *Artificial Intelligence*, 170(11):983–1016, 2006.

- [5] M. C. Cooper, S. de Givry, M. Sanchez, T. Schiex, M. Zytnicki, and T. Werner. Soft arc consistency revisited. *Artif. Intell.*, 174(7-8):449–478, May 2010.
- [6] V. Dalmau and J. Pearson. Closure functions and width 1 problems. In *In CP 1999*, pages 159–173. Springer-Verlag, 1999.
- [7] T. Feder and M. Vardi. The computational structure of monotone monadic SNP a and constraint satisfaction: A study through Datalog and group theory. *SIAM Journal on Computing*, 28(1):57–104, 1998.
- [8] V. Kolmogorov. Convergent tree-reweighted messages passing. *PAMI*, 28(10):1568–1583, 2006.
- [9] V. Kolmogorov and T. Schoenemann. Generalized sequential tree-reweighted message passing. *CoRR*, abs/1205.6352, 2012.
- [10] V. Kolmogorov and S. Živný. The complexity of conservative valued CSPs. In *SODA*, 2012.
- [11] A. Koster, C. P. M. van Hoesel, and A. W. J. Kolen. The partial constraint satisfaction problem: Facets and lifting theorems. *Operation Research Letters*, 23(3-5):89–97, 1998.
- [12] Gabor Kun, Ryan O’Donnell, Suguru Tamaki, Yuichi Yoshida, and Yuan Zhou. Linear programming, width-1 CSPs, and robust satisfaction. In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*, ITCS ’12, pages 484–495, New York, NY, USA, 2012. ACM.
- [13] O. Meshi and A. Globerson. An alternating direction method for dual map lp relaxation. In *European Conference on Machine Learning*, 2011.
- [14] Prasad Raghavendra. Approximating NP-hard Problems: Efficient Algorithms and their Limits. *PhD Thesis*, 2009.
- [15] M. I. Schlesinger. Syntactic analysis of two-dimensional visual signals in noisy conditions. *Kibernetika*, 4:113–130, 1976. In Russian.
- [16] David Sontag, Amir Globerson, and Tommi Jaakkola. Introduction to dual decomposition for inference. In Suvrit Sra, Sebastian Nowozin, and Stephen J. Wright, editors, *Optimization for Machine Learning*. MIT Press, 2011.
- [17] J. Thapper and S. Živný. The power of linear programming for valued CSPs. In *FOCS*, 2012.
- [18] M. J. Wainwright, T. S. Jaakkola, and A. S. Willsky. MAP estimation via agreement on (hyper)trees: Message-passing and linear-programming approaches. *IEEE Transactions on Information Theory*, 51(11):3697–3717, November 2005.
- [19] T. Werner. A linear programming approach to max-sum problem: A review. *PAMI*, 29(7):1165–1179, 2007.